MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

[illegible]on: Mon Aug 3 11:07:05 1998; MasPar time 1670.08 Seconds
1477.192 Million cell updates/sec
Tabular output not generated.

Title: >US-08-915-659A-6
Description: (1-1343) from US08915659A.seq
Perfect Score: 1343
N.A. Sequence: 1 CTGTAGCAGGCAGAGCTTAC.....TTAATAAACTCACGGAATTC 1343
Comp: GACATCGTCCGTCTCGAATG.....AATTATTTGAGTGCCTTAAG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532259 seqs, 918475165 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl54
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
Database: genbank106
12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_vi

[illegible]istics: Mean 11.214; Variance 4.982; scale 2.251

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	845	62.9	868	21	AB009849	Homo sapiens mRNA for	0.00e+00
2	423	31.5	1322	22	MUSNEU	Mouse mRNA for neurops	0.00e+00
3	261	19.4	2711	21	AB012761	Homo sapiens HNP gene	7.74e-207
4	77	5.7	849	22	MMNGFB	Mouse mRNA for gamma s	9.16e-40
5	74	5.5	428	22	MMTAM1	Mouse mRNA for gamma-7	3.02e-37
6	72	5.4	1454	21	AF024605	Homo sapiens serine pr	1.41e-35
7	71	5.3	689	22	RNTKALRC	Rat tissue mRNA for ka	9.54e-35
8	71	5.3	867	22	RATPKLK	rat pancreatic preprok	9.54e-35
9	71	5.3	873	22	RATKALPS	Rat submaxillary gland	9.54e-35
10	71	5.3	1419	21	D78203	Human mRNA for neurosi	9.54e-35
11	71	5.3	1451	21	AF013988	Homo sapiens serine pr	9.54e-35
12	71	5.3	1457	21	S82666	NES1=normal epithelial	9.54e-35
13	71	5.3	1506	21	HSU62801	Human protease M mRNA,	9.54e-35
14	70	5.2	762	22	RATKALLB	Rattus norvegicus (clo	6.44e-34
15	69	5.1	518	22	RATKALSA	Rat submaxillary gland	4.33e-33
16	69	5.1	870	22	MUSEGFBPM	Mouse major epidermal	4.33e-33
17	67	5.0	723	22	RATKALLA	Rattus norvegicus (clo	1.92e-31
18	65	4.8	800	20	BABKLK1SP	Papio hamadryas tissue	8.37e-30
19	65	4.8	830	22	MUSNGF7S	Mouse nerve growth fac	8.37e-30
c 20	65	4.8	7218	17	I66494	Sequence 14 from paten	8.37e-30
21	63	4.7	858	20	MACKAL	Macaca fascicularis (c	3.56e-28
22	63	4.7	925	17	I09127	Sequence 2 from Patent	3.56e-28
23	62	4.6	1414	20	HSRNASPH	H.sapiens ACO mRNA for	2.30e-27
24	60	4.5	659	22	RATRGK8A3	Rat kallikrein (rGK-8)	9.44e-26
25	60	4.5	832	15	CFKALLIK	C.familiaris dKlk-2 mR	9.44e-26
26	60	4.5	863	22	MUSMGK6	Mouse mRNA for tissue	9.44e-26
27	61	4.5	2448	22	RATRSKG	Rat kallikrein-like se	1.48e-26
28	61	4.5	5574	21	AF055481	Homo sapiens normal ep	1.48e-26
29	59	4.4	1110	22	AF016269	Rattus norvegicus myel	5.99e-25
30	57	4.2	738	17	I02159	Sequence 7 from Patent	2.36e-23
31	57	4.2	738	17	I08522	Sequence 2 from Patent	2.36e-23
32	57	4.2	738	20	HUMKALA	Human kidney kallikrei	2.36e-23
33	57	4.2	819	15	DOGTRYPA	Dog pancreatic anionic	2.36e-23
34	57	4.2	871	20	HUMKALX	Human kallikrein mRNA,	2.36e-23
35	57	4.2	871	17	E01205	DNA sequence of human	2.36e-23
36	56	4.2	1541	21	S39329	glandular kallikrein-1	1.46e-22
37	57	4.2	6257	22	RATKALA	Rat renal kallikrein g	2.36e-23
38	55	4.1	871	20	HSKALLI	Human mRNA for preprok	9.03e-22
39	55	4.1	1463	22	MUSGFBPB3	Mouse epidermal growth	9.03e-22
40	55	4.1	9433	22	MUSKAL07	Mouse pseudo-kallikrei	9.03e-22
41	55	4.1	9433	22	MMKALL	Mouse complete gene fo	9.03e-22
42	54	4.0	500	22	MUSGKAL	Mouse glandular kallik	5.52e-21
43	54	4.0	992	17	I67863	Sequence 13 from paten	5.52e-21
44	54	4.0	1140	15	SSU76256	Sus scrofa enamel matr	5.52e-21
45	54	4.0	1446	20	HUMAPS	Homo sapiens prostate-	5.52e-21

ALIGNMENTS

RESULT 1
LOCUS AB009849 868 bp mRNA PRI 03-JUN-1998
DEFINITION Homo sapiens mRNA for neuropsin, complete cds.
ACCESSION AB009849
NID g3176386
KEYWORDS neuropsin.
SOURCE Homo sapiens hippocampus cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Yoshida,S., Kitayoshi,H., Kuwae,K., Inoue,N. and Shiosaka,S.
TITLE Sequence analysis and expression of human neuropsin cDNA and gene
JOURNAL Gene 213, 9-16 (1998)
REFERENCE 2 (bases 1 to 868)
AUTHORS Yoshida,S.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) to the DDBJ/EMBL/GenBank databases.
Shigetaka Yoshida, Nara Institute of Science and Technology,
Division of Structural Cell Biology; Takayama 8916-5, Ikoma, Nara
630-0101, Japan (E-mail:syoshida@bs.aist-nara.ac.jp,
Tel:81-743-72-5411, Fax:81-743-72-5419)
COMMENT Sequence updated (01-Jun-1998).
FEATURES Location/Qualifiers
source 1. .868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hippocampus"
sig_peptide 35. .118
CDS 35. .817
/codon_start=1
/product="neuropsin"
/db_xref="PID:d1029613"
/db_xref="PID:g3176387"
/translation="MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQ
PWQAALFQGQQLLCGGVLVGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQ
SIPHPCYNSSDVEDHNHDLMLLQLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTV
TSPRENFPDTLNCAEVKIFPQKKCEDAYPGQITDGMVCAGSSKGADTCQGDSGGPLVC
DGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSKG"

* Check IFN!

```
     mat_peptide     131. .814
                     /product="neuropsin"
BASE COUNT      206 a    253 c    242 g    167 t
ORIGIN

  Query Match           62.9%;  Score 845;  DB 21;  Length 868;
  Best Local Similarity 100.0%;  Pred. No. 0.00e+00;
  Matches   845;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     11 CCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGTGCGGCCAAGACGTGG 70
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    492 CCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGTGCGGCCAAGACGTGG 551

Db     71 ATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAGGAGGACAAG 130
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    552 ATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAGGAGGACAAG 611

Db    131 GTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTGTTCCAG 190
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    612 GTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTGTTCCAG 671

Db    191 GGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCTGCC 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    672 GGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCTGCC 731

Db    251 CACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT 310
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    732 CACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT 791

Db    311 GGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGC 370
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    792 GGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGC 851

Db    371 GATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTG 430
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    852 GATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTG 911

Db    431 GGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGC 490
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    912 GGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGC 971

Db    491 ACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAAC 550
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    972 ACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAAC 1031

Db    551 TGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATC 610
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1032 TGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATC 1091

Db    611 ACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCT 670
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1092 ACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCT 1151

Db    671 GGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCC 730
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1152 GGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCC 1211

Db    731 TGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATC 790
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1212 TGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATC 1271

Db    791 AAGAAGATCATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAA 850
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1272 AAGAAGATCATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAA 1331

Db    851 CTCAC 855
        |||||
Qy   1332 CTCAC 1336


RESULT    2
LOCUS       MUSNEU       1322 bp    mRNA            ROD       09-DEC-1996
DEFINITION  Mouse mRNA for neuropsin, complete cds.
ACCESSION   D30785
NID         g1648847
KEYWORDS    Neuropsin; serine protease.
SOURCE      Mus musculus (strain:BALB/c) hippocampus neuron cDNA to mRNA,
            clone_lib:lambda-gt10 clone:NP.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 1322)
  AUTHORS   Chen,Z.-L., Yoshida,S., Kato,K., Momota,Y., Suzuki,J., Tanaka,T.,
            Ito,J., Nishino,H., Aimoto,S., Kiyama,H. and Shiosaka,S.
  TITLE     Expression and activity-dependent changes of a novel limbic-serine
            protease gene in the hippocampus
  JOURNAL   J. Neurosci. 15 (7 Pt 2), 5088-5097 (1995)
  MEDLINE   95348817
REFERENCE   2  (bases 1 to 1322)
  AUTHORS   Shiosaka,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1994) to the DDBJ/EMBL/GenBank databases. Sadao
            Shiosaka, Nara Institute of Science and Technology, Department of
            Structural Cellular Biology; 8916-5 Takayama, Ikoma, Nara 630-01,
            Japan (E-mail:sshiosak@bx.aist-nara.ac.jp, Tel:07437-2-5410,
            Fax:07437-2-5419)
FEATURES             Location/Qualifiers
     source          1. .1322
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /cell_type="neuron"
                     /clone="NP"
                     /clone_lib="lambda-gt10"
                     /tissue_type="hippocampus"
     misc_feature    482. .490
                     /note="kozak sequence"
     CDS             487. .1269
                     /note="nervous system"
                     /codon_start=1
                     /product="neuropsin"
                     /db_xref="PID:d1007022"
                     /db_xref="PID:g1020091"
                     /translation="MGRPPPCAIQPWILLLLFMGAWAGLTRAQGSKILEGRECIPHSQ
                     PWQAALFQGERLICGGVLVGDRWVLTAAHCKKQKYSVRLGDHSLQSRDQPEQEIQVAQ
                     SIQHPCYNNSNPEDHSHDIMLIRLQNSANLGDKVKPVQLANLCPKVGQKCIISGWGTV
                     TSPQENFPNTLNCAEVKIYSQNKCERAYPGKITEGMVCAGSSNGADTCQGDSGGPLVC
                     DGMLQGITSWGSDPCGKPEKPGVYTKICRYTTWIKKTMDNRD"
     polyA_signal    1300. .1305
     polyA_site      1322
BASE COUNT      309 a    414 c    339 g    260 t
ORIGIN

  Query Match           31.5%;  Score 423;  DB 22;  Length 1322;
  Best Local Similarity 76.7%;  Pred. No. 0.00e+00;
  Matches   607;  Conservative    0;  Mismatches 184;  Indels   0;  Gaps   0;

Db    483 CACCATGGGACGCCCCCCACCCTGTGCAATCCAGCCGTGGATCCTTCTGCTTCTGTTCAT 542
        |||||||||||||||| ||| |||| | || ||||||| | ||||  || |
Qy    512 CACCATGGGACGCCCCCGACCTCGTGCGGCCAAGACGTGGATGTTCCTGCTCTTGCTGGG 571

Db    543 GGGAGCGTGGGCAGGGCTCACCAGAGCTCAGGGCTCCAAGATCCTGGAAGGTCGAGAGTG 602
        |||||| |||||||| | | |||| || ||||    |||| | ||||  |||| |||||
Qy    572 GGGAGCCTGGGCAGGACACTCCAGGGCACAGGAGGACAAGGTGCTGGGGGGTCATGAGTG 631

Db    603 TATACCCCACTCCCAGCCTTGGCAGGCAGCCTTGTTCCAGGGCGAGAGACTGATCTGTGG 662
          |||||| || ||||||||||||||| |||||||||||||||| ||  |||  |||||||
Qy    632 CCAACCCCATTCGCAGCCTTGGCAGGCGGCCTTGTTCCAGGGCCAGCAACTACTCTGTGG 691

Db    663 GGGTGTCCTGGTTGGAGACAGATGGGTCCTCACGGCAGCCCACTGCAAAAAACAGAAGTA 722
          |||||||| || || | ||  |||||||| || || |||||||| ||||||| ||| ||
Qy    692 CGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCTGCCCACTGTAAAAAACCGAAATA 751

Db    723 CTCCGTGCGTCTGGGTGATCATAGCCTCCAGAGCAGAGATCAGCCGGAGCAGGAGATCCA 782
          | | || || ||||| || || ||||| |||| | ||||   || ||||| || || |
```

```
Qy    752 CACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGATGGCCCAGAGCAAGAAATACC 811

Db    783 GGTGGCTCAGTCTATCCAGCATCCTTGCTACAACAACAGCAACCCAGAAGATCACAGTCA 842

Qy    812 TGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGATGTGGAGGACCACAACCA 871

Db    843 CGATATAATGCTCATTCGACTGCAGAACTCAGCAAACCTCGGGGACAAGGTGAAGCCGGT 902

Qy    872 TGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCCAAAGTGAAGCCCAT 931

Db    903 CCAACTGGCCAATCTGTGTCCCAAAGTTGGCCAGAAGTGCATCATATCAGGCTGGGGCAC 962

Qy    932 CAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCAC 991

Db    963 TGTCACCAGCCCTCAAGAGAACTTTCCAAACACCCTCAACTGTGCGGAAGTGAAAATCTA 1022

Qy    992 TGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTT 1051

Db   1023 TTCCCAGAACAAGTGTGAGAGAGCCTATCCAGGGAAGATCACCGAGGGCATGGTCTGTGC 1082

     1052 TCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGC 1111

Db   1083 TGGCAGCAGCAATGGAGCTGACACGTGCCAGGGTGACTCAGGAGGCCCTCTGGTGTGCGA 1142

Qy   1112 AGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGA 1171

Db   1143 CGGGATGCTCCAGGGCATCACCTCATGGGGCTCAGACCCCTGTGGGAAACCCGAGAAACC 1202

Qy   1172 TGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACC 1231

Db   1203 TGGAGTCTACACCAAAATCTGCCGCTACACTACCTGGATCAAGAAGACCATGGACAACAG 1262

Qy   1232 TGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAA 1291

Db   1263 GGACTGATCCT 1273

Qy   1292 GGGCTGATTCT 1302


RESULT    3
LOCUS       AB012761     2711 bp    DNA             PRI       03-JUN-1998
DEFINITION  Homo sapiens HNP gene for neuropsin, partial cds.
ACCESSION   AB012761
NID         g3176392
KEYWORDS    neuropsin.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Yoshida,S., Kitayoshi,H., Kuwae,K., Inoue,N. and Shiosaka,S.
  TITLE     Sequence analysis and expression of human neuropsin cDNA and gene
  JOURNAL   Gene 213, 9-16 (1998)
REFERENCE   2  (bases 1 to 2711)
  AUTHORS   Yoshida,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-1998) to the DDBJ/EMBL/GenBank databases.
            Shigetaka Yoshida, Nara Institute of Science and Technology,
            Division of Structural Cell Biology; Takayama 8916-5, Ikoma, Nara
            630-0101, Japan (E-mail:syoshida@bs.aist-nara.ac.jp,
            Tel:81-743-72-5411, Fax:81-743-72-5419)
FEATURES             Location/Qualifiers
     source          1. .2711
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CAAT_signal     103. .107
     exon            <1100. .1125
     intron          1126. .1479
     exon            1480. .1557
     gene            1488. .2657
                     /gene="HNP"
     CDS             join(1488. .1557,2070. .2229,2395. .>2657)
                     /gene="HNP"
                     /codon_start=1
                     /product="neuropsin"
                     /db_xref="PID:d1029616"
                     /db_xref="PID:g3176393"
                     /translation="MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQ
                     PWQAALFQGQQLLCGGVLVGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQ
                     SIPHPCYNSSDVEDHNHDLMLLQLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTV
                     TSPR"
     sig_peptide     join(1488. .1557,2070. .2071)
                     /gene="HNP"
     intron          1558. .2069
                     /gene="HNP"
     exon            2070. .2229
                     /gene="HNP"
     intron          2230. .2394
                     /gene="HNP"
     exon            2395. .2657
                     /gene="HNP"
     intron          2658. .>2711
BASE COUNT      554 a    840 c    756 g    561 t
ORIGIN

  Query Match           19.4%;  Score 261;  DB 21;  Length 2711;
  Best Local Similarity  99.6%;  Pred. No. 7.74e-207;
  Matches   262;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Db   2395 GAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGATGGCCCAGAGCAAGA 2454

Qy    746 GAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGATGGCCCAGAGCAAGA 805

Db   2455 AATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGATGTGGAGGACCA 2514

Qy    806 AATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGATGTGGAGGACCA 865

Db   2515 CAACCATGATCTGATGCTTCTTCAATTGCGTGACCAGGCATCCCTGGGGTCCAAAGTGAA 2574

Qy    866 CAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCCAAAGTGAA 925

Db   2575 GCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTG 2634

Qy    926 GCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTG 985

Db   2635 GGGCACTGTCACCAGTCCCCGAG 2657

Qy    986 GGGCACTGTCACCAGTCCCCGAG 1008


RESULT    4
LOCUS       MMNGFB       849 bp    RNA             ROD       30-MAR-1995
DEFINITION  Mouse mRNA for gamma subunit of nerve growth factor (NGF).
ACCESSION   X01389
NID         g53373
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 849)
  AUTHORS   Ullrich,A., Gray,A., Wood,W.I., Hayflick,J. and Seeburg,P.H.
  TITLE     Isolation of a cDNA clone coding for the gamma-subunit of mouse
            nerve growth factor using a high-stringency selection procedure
  JOURNAL   DNA 3 (5), 387-392 (1984)
  MEDLINE   85076169
FEATURES             Location/Qualifiers
     source          1. .849
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             18. .803
                     /note="gamma NGF precursor"
                     /codon_start=1
                     /db_xref="PID:g53374"
                     /db_xref="SWISS-PROT:P00756"
```

```
                     /translation="MWFLILFLALSLGGIDAAPPVQSRIVGGFKCEKNSQPWHVAVYR
                     YTQYLCGGVLLDPNWVLTAAHCYDDNYKVWLGKNNLFKDEPSAQHRFVSKAIPHPGFN
                     MSLMRKHIRFLEYDYSNDLMLLRLSKPADITDTVKPITLPTEEPKLGSTCLASGWGSI
                     TPTKFQFTDDLYCVNLKLLPNEDCAKAHIEKVTDAMLCAGEMDGGKDTCKGDSGGPLI
                     CDGVLQGITSWGHTPCGEPDMPGVYTKLNKFTSWIKDTMAKNP"
     sig_peptide     18. .89
                     /note="put. signal peptide"
     mat_peptide     90. .800
                     /note="gamma NGF"
     misc_feature    90. .388
                     /note="proteolytic peptide B1"
     misc_feature    389. .400
                     /note="connecting peptide"
     misc_feature    401. .800
                     /note="proteolytic peptide C2"
     misc_feature    825. .830
                     /note="polyA signal"
     polyA_site      849
                     /note="polyA site"
BASE COUNT      227 a    236 c    196 g    190 t
ORIGIN

  Query Match             5.7%;  Score 77;  DB 22;  Length 849;
  Best Local Similarity  61.6%;  Pred. No. 9.16e-40;
  Matches   244;  Conservative    0;  Mismatches 149;  Indels   3;  Gaps   2;

Db     356 TGAGTATGACTACAGCAATGACCTGATGCTGCTCCGCCTCAGCAAGCCTGCTGACATCAC 415
           || | | ||| ||| | |||| |||||||| || |  ||  |  | |  ||   | |
Qy     854 TGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGG 913

Db     416 AGATACTGTGAAGCCCATCACCCTGCCCACTGAGGAGCCCAAGCTGGGGAGCACATGCCT 475
           |  |||||||||||||| |||| |   | |     || |||  ||    |  |||
Qy     914 GTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCAC 973

Db     476 AGCCTCAGGCTGGGGCAGCATTACACCCACCAAATTCCAATTCACAGATGATCTCTACTG 535
           | ||||||||||||||   | ||     ||  |    | ||  | ||   |||| ||||
Qy     974 CGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTG 1033

Db     536 TGTGAACCTCAAGCTCCTGCCTAATGAGGACTGTGCCAAAGCCCACATAGAGAAGGTGAC 595
           ||   |  | ||  || | ||  |  || | ||||   | ||  ||   | | || | ||
Qy    1034 TGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCAC 1093

Db     596 AGATGCCATGCTGTGTGCAGGAGAGATGGATGGAGGCAAAGACACTTGCAAGGGTGACTC 655
           |||||  |||| | ||||||||  ||  | |  | |||   |||||  ||| |||| || ||
Qy    1094 AGATGGCATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTC 1150

Db     656 AGGAGGCCCACTGATCTGTGATGGTGTTCTCCAAGGTATCACATCATGGGGCCATACCCC 715
           ||||||||  |||  | ||||||||||  |||||  ||  ||||||||  ||||||     |||
Qy    1151 TGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCC 1210

Db     716 ATGCGGTGAACCTGATATGCCAGGCGTCTACACCAA 751
           || ||      | || |  || |||||||| |||||
Qy    1211 CTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT    5
LOCUS       MMTAM1        428 bp    RNA             ROD       10-JUL-1995
DEFINITION  Mouse mRNA for gamma-7S nerve growth factor (y-NGF) fragment.
ACCESSION   X00472
NID         g54260
KEYWORDS    complementary DNA; nerve growth factor; serine protease.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE   1  (bases 1 to 428)
  AUTHORS   Howles,P.N., Dickinson,D.P., DiCaprio,L.L., Woodworth-Gutai,M. and
            Gross,K.W.
  TITLE     Use of a cDNA recombinant for the gamma-subunit of mouse nerve
            growth factor to localize members of this multigene family near the
            TAM-1 locus on chromosome 7
  JOURNAL   Nucleic Acids Res. 12 (6), 2791-2805 (1984)
  MEDLINE   84169573
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             <1. .408
                     /note="gamma-NGF"
                     /codon_start=1
                     /db_xref="PID:g54261"
                     /translation="SKPADITDTVKPITLPTEEPKLGSTCLASGWGSITPTKFQFTDD
                     LYCVNLKLLPNEDCAKAHIEKVTDAMLCAGEMEGGKDTCKGDSGGPLICDGVLQGITS
                     WGHTPCGEPDMPGVYTKLNKFTSWIKDTMAKNP"
BASE COUNT      123 a    116 c    102 g     87 t
ORIGIN

  Query Match             5.5%;  Score 74;  DB 22;  Length 428;
  Best Local Similarity  63.5%;  Pred. No. 3.02e-37;
  Matches   209;  Conservative    0;  Mismatches 117;  Indels   3;  Gaps   2;

Db      28 GTGAAGCCCATCACCCTGCCCACTGAGGAGCCCAAGCTGGGGAGCACATGCCTAGCCTCA 87
           |||||||||||||| |||| |   | |     || |||  ||    |  |||   | ||||
Qy     921 GTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCA 980

Db      88 GGCTGGGGCAGCATTACACCCACCAAATTCCAATTCACAGATGATCTCTACTGTGTGAAC 147
           ||||||||||   | ||     ||  |    | ||  | ||   |||| ||||||   |
Qy     981 GGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAGAA 1040

Db     148 CTCAAGCTCCTGCCTAATGAGGACTGTGCCAAAGCCCACATAGAGAAGGTGACAGATGCC 207
           | ||  || | ||  |  || | ||||   | ||  ||   | | || | ||||||| |
Qy    1041 GTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGC 1100

Db     208 ATGCTGTGTGCAGGAGAGATGGAAGGAGGCAAAGACACTTGCAAGGGTGACTCAGGAGGC 267
           ||| | ||||||||  ||  | || | |||   ||||| ||| |||| || || ||||||
Qy    1101 ATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTCTGGAGGC 1157

Db     268 CCACTGATCTGTGATGGTGTTCTCCAAGGTATCACATCATGGGGCCATACCCCATGCGGT 327
           || ||| | ||||||||||  ||||| || ||||||||  ||||||     ||| || ||
Qy    1158 CCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGG 1217

Db     328 GAACCTGATATGCCAGGCGTCTACACCAA 356
           | || |  || |||||||| |||||
Qy    1218 AGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT    6
LOCUS       AF024605     1454 bp    mRNA            PRI       24-OCT-1997
DEFINITION  Homo sapiens serine protease-like protease (nesl) mRNA, complete
            cds.
ACCESSION   AF024605 S82666
NID         g2558911
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1454)
  AUTHORS   Liu,X.L., Wazer,D.E., Watanabe,K. and Band,V.
  TITLE     Identification of a novel serine protease-like gene, the expression
            of which is down-regulated during breast cancer progression
  JOURNAL   Cancer Res. 56 (14), 3371-3379 (1996)
  MEDLINE   96320486
REFERENCE   2  (bases 1 to 1454)
  AUTHORS   Liu,X.-L., Wazer,D.E., Watanabe,K. and Band,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1997) Radiation Oncology, New England Medical
            Center Hospital, Tufts University School of Medicine, NEMC #824,
            750 Washington St., Boston, MA 02111, USA
FEATURES             Location/Qualifiers
     source          1. .1454
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="76N"
                     /cell_type="mammary epithelial"
```

```
     gene            1. .1454
                     /gene="nes1"
     CDS             82. .912
                     /gene="nes1"
                     /note="NES1"
                     /codon_start=1
                     /product="serine protease-like protease"
                     /db_xref="PID:g2558912"
                     /translation="MRAPHLHLSAASGARALAKLLPLLMAQLWAAEAALLPQNDTRLD
                     PEAYGAPCARGSQPWQVSLFNGLSFHCAGVLVDQSWVLTAAHCGNKPLWARVGDDHLL
                     LLQGEQLRRTTRSVVHPKYHQGSGPILPRRTDEHDLMLLKLARPVVPGPRVRALQLPY
                     RCAQPGDQCQVAGWGTTAARRVKYNKGLTCSSITILSPKECEVFYPGVVTNNMICAGL
                     DRGQDPCQSDSGGPLVCDETLQGILSWGVYPCGSAQHPAVYTQICKYMSWINKVIRSN
                     "
BASE COUNT      289 a    481 c    377 g    307 t
ORIGIN

  Query Match           5.4%;  Score 72;  DB 21;  Length 1454;
  Best Local Similarity  60.7%;  Pred. No. 1.41e-35;
  Matches   205;  Conservative    0;  Mismatches 133;  Indels   0;  Gaps   0;

Db     571 CAGCCCGGAGACCAGTGCCAGGTTGCTGGCTGGGGCACCACGGCCGCCCGGAGAGTGAAG 630
           ||||| ||  |  |||||   ||  | ||||||||||     ||   |   ||| |||
Qy     954 CAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAAT 1013

Db     631 TACAACAAGGGCCTGACCTGCTCCAGCATCACTATCCTGAGCCCTAAAGAGTGTGAGGTC 690
           |     |    || | |||  |     | |  ||| |   ||  ||  |||||||||
Qy    1014 TTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGAT 1073

Db     691 TTCTACCCTGGCGTGGTCACCAACAACATGATATGTGCTGGACTGGACCGGGGCCAGGAC 750
           |||||  ||   | ||||  |   |||| | ||||| ||       |   ||    |||
Qy    1074 GCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGAC 1133

Db     751 CCTTGCCAGAGTGACTCTGGAGGCCCCCTGGTCTGTGACGAGACCCTCCAAGGCATCCTC 810
           | ||||||  | ||  |||||||||||||||||  |||||  |   | |||||  ||||||
Qy    1134 ACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACA 1193

Db     811 TCGTGGGGTGTTTACCCCTGTGGCTCTGCCCAGCATCCAGCTGTCTACACCCAGATCTGC 870
           || |||||     ||||||||||     || |  | || |  |||||  ||| | ||||||
Qy    1194 TCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATCTGC 1253

Db     871 AAATACATGTCCTGGATCAATAAAGTCATACGCTCCAA 908
           |||  ||  |||||||||  ||  |||||  ||  |||
Qy    1254 CGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAA 1291


RESULT    7
LOCUS       RNTKALRC      689 bp    RNA             ROD       13-JUN-1995
DEFINITION  Rat tissue mRNA for kallikrein C-terminal region (EC 3.4.21.8).
ACCESSION   X03560
NID         g57370
KEYWORDS    kallikrein; protease; serine protease.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 189 to 689)
  AUTHORS   Gerald,W.L., Chao,J. and Chao,L.
  TITLE     Immunological identification of rat tissue kallikrein cDNA and
            characterization of the kallikrein gene family
  JOURNAL   Biochim. Biophys. Acta 866 (1), 1-14 (1986)
  MEDLINE   86131678
REFERENCE   2  (bases 1 to 188)
  AUTHORS   Chao,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1986) Medical University of South Carolina, 171
            Ashley Avenue, Charleston, South Carolina 29425-2211
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             <1. .646
                     /codon_start=2
                     /product="kallikrein"
                     /db_xref="PID:g818030"
                     /translation="YLCGGVLIDPSWVITAAHCATDNYQVWLGRNNLYEDEPFAQHRL
                     VSQSFPHPGFNQDLIWNHTRQPGDDYSNDLMLLHLSQPADITDGVKVIDLPIEEPKVG
                     STCLASGWGSITPDGLELSDDLQCVNIDLLSNEKCVEAHKEEVTDLMLCAGEMDGGKD
                     TCKGDSGGPLICNGVLQGITSWGFNPCGEPKKPGIYTKLIKFTPWIKEVMKENP"
     polyA_signal    670. .675
                     /note="putative"
     polyA_site      689
BASE COUNT      181 a    189 c    178 g    141 t
ORIGIN

  Query Match           5.3%;  Score 71;  DB 22;  Length 689;
  Best Local Similarity  60.9%;  Pred. No. 9.54e-35;
  Matches   241;  Conservative    0;  Mismatches 152;  Indels   3;  Gaps   2;

Db     199 TGGGGACGACTACAGCAATGATTTGATGCTGCTCCACCTCAGCCAGCCTGCGGACATCAC 258
           || ||| ||| ||| | ||||| ||||||| || || ||  |  | |  ||   | |
Qy     854 TGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGG 913

Db     259 AGATGGTGTGAAGGTCATCGATCTGCCCATTGAGGAGCCCAAGGTGGGGAGCACCTGCCT 318
               ||||||  ||||   ||| |   | |     || ||   ||    |  |||
Qy     914 GTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCAC 973

Db     319 TGCCTCGGGCTGGGGCAGCATCACACCTGACGGATTGGAATTAAGTGATGATCTCCAGTG 378
           | ||| ||||||||||   ||||   |  | ||  | | ||   |||   |||| | ||
Qy     974 CGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTG 1033

Db     379 TGTGAACATCGATCTTCTGTCTAATGAGAAGTGCGTCGAGGCACACAAAGAAGAGGTGAC 438
           ||   |  |  |  |  |  |  |  |  |||||||  |  || ||  ||   |  || | ||
Qy    1034 TGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCAC 1093

Db     439 AGATCTCATGCTGTGTGCAGGAGAGATGGATGGGGGCAAAGACACTTGCAAGGGTGACTC 498
           ||||  |||| | ||||||||  ||  | |  |||||   ||||| ||| |||| || ||
Qy    1094 AGATGGCATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTC 1150

Db     499 AGGAGGCCCCCTCATCTGTAATGGTGTGCTCCAAGGCATCACGTCCTGGGGCTTTAACCC 558
           ||||||||||  | ||| ||||||  |||||  |||||||| ||||||||||   ||||
Qy    1151 TGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCC 1210

Db     559 ATGCGGTGAACCCAAGAAGCCAGGCATCTACACCAA 594
           || ||     || | || || ||| |||| |||||
Qy    1211 CTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT    8
LOCUS       RATPKLK       867 bp    mRNA            ROD       02-MAY-1983
DEFINITION  rat pancreatic preprokallikrein mrna.
ACCESSION   J00758
NID         g206200
KEYWORDS    kallikrein; protease; serine protease.
SOURCE      rat pancreatic tissue.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 1 to 867)
  AUTHORS   Swift,G.H., Dagorn,J.C., Ashley,P.L., Cummings,S.W. and
            MacDonald,R.J.
  TITLE     rat pancreatic kallikrein mrna: nucleotide sequence and amino acid
            sequence of the encoded preproenzyme
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 79, 7263-7267 (1982)
  MEDLINE   83117659
FEATURES             Location/Qualifiers
     source          1. .867
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
BASE COUNT      212 a    242 c    224 g    189 t
ORIGIN

  Query Match           5.3%;  Score 71;  DB 22;  Length 867;
  Best Local Similarity  60.9%;  Pred. No. 9.54e-35;
```

```
  Matches   241;  Conservative    0;  Mismatches 152;  Indels   3;  Gaps   2;

Db     375 TGGGGACGACTACAGCAATGATTTGATGCTGCTCCACCTCAGCCAGCCTGCGGACATCAC 434
           || ||| ||| ||| | ||||| ||||||| || || ||  |  | |  ||   | |
Qy     854 TGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGG 913

Db     435 AGATGGTGTGAAGGTCATCGATCTGCCCATTGAGGAGCCCAAGGTGGGGAGCACCTGCCT 494
             ||||||  ||||   ||| |   | |     || ||   ||    |  |||
Qy     914 GTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCAC 973

Db     495 TGCCTCGGGCTGGGGCAGCATCACACCTGACGGATTGGAATTAAGTGATGATCTCCAGTG 554
           | ||| ||||||||||   ||||   |  | ||  | | ||   |||   |||| | ||
Qy     974 CGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTG 1033

Db     555 TGTGAACATCGATCTTCTGTCTAATGAGAAGTGCGTCGAGGCACACAAAGAAGAGGTGAC 614
           ||   |  |  |  |  |  |  | |||||||  |  || ||  ||   |   || | ||
Qy    1034 TGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCAC 1093

Db     615 AGATCTCATGCTGTGTGCAGGAGAGATGGATGGGGGCAAAGACACTTGCAAGGGTGACTC 674
           ||||  |||| | ||||||||  ||  | |  |||||   ||||| ||| |||| || ||
Qy    1094 AGATGGCATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTC 1150

Db     675 AGGAGGCCCCCTCATCTGTAATGGTGTGCTCCAAGGCATCACGTCCTGGGGCTTTAACCC 734
           ||||||||||  | ||| ||||||  ||||| ||||||| ||||||||||   ||||
Qy    1151 TGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCC 1210

Db     735 ATGCGGTGAACCCAAGAAGCCAGGCATCTACACCAA 770
           || ||     || | || || ||| |||| |||||
Qy    1211 CTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT    9
LOCUS       RATKALPS      873 bp    mRNA            ROD       01-SEP-1988
DEFINITION  Rat submaxillary gland PS kallikrein mRNA, complete cds.
ACCESSION   M11563
NID         g205029
KEYWORDS    .
SOURCE      Rat submaxillary gland, cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 1 to 873)
  AUTHORS   Ashley,P.L. and MacDonald,R.J.
  TITLE     Kallikrein-related mRNAs of the rat submaxillary gland: Nucleotide
            sequences of four distinct types including tonin
  JOURNAL   Biochemistry 24, 4512-4520 (1985)
  MEDLINE   86051477
FEATURES             Location/Qualifiers
     source          1. .873
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             31. .828
                     /note="PS kallikrein"
                     /codon_start=1
                     /db_xref="PID:g205030"
                     /translation="MPVTMWFLILFLALSLGRNDAAPPVQSRVVGGYNCEMNSQPWQV
                     AVYYFGEYLCGGVLIDPSWVITAAHCATDNYQVWLGRNNLYEDEPFAQHRLVSQSFPH
                     PGFNQDLIWNHTRQPGDDYSNDLMLLHLSQPADITDGVKVIDLPIEEPKVGSTCLASG
                     WGSITPDGLELSDDLQCVNIDLLSNEKCVEAHKEEVTDLMLCAGEMDGGKDTCKGDSG
                     GPLICNGVLQGITSWGFNPCGEPKKPGIYTKLIKFTPWIKEVMKENP"
BASE COUNT      213 a    245 c    225 g    190 t
ORIGIN

  Query Match             5.3%;  Score 71;  DB 22;  Length 873;
  Best Local Similarity  60.9%;  Pred. No. 9.54e-35;
  Matches   241;  Conservative    0;  Mismatches 152;  Indels   3;  Gaps   2;

Db     381 TGGGGACGACTACAGCAATGATTTGATGCTGCTCCACCTCAGCCAGCCTGCGGACATCAC 440
           || ||| ||| ||| | ||||| ||||||| || || ||  |  | |  ||   | |
Qy     854 TGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGG 913

Db     441 AGATGGTGTGAAGGTCATCGATCTGCCCATTGAGGAGCCCAAGGTGGGGAGCACCTGCCT 500
             ||||||  ||||   ||| |   | |     || ||   ||    |  |||
Qy     914 GTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCAC 973

Db     501 TGCCTCGGGCTGGGGCAGCATCACACCTGACGGATTGGAATTAAGTGATGATCTCCAGTG 560
           | ||| ||||||||||   ||||   |  | ||  | | ||   |||   |||| | ||
Qy     974 CGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTG 1033

Db     561 TGTGAACATCGATCTTCTGTCTAATGAGAAGTGCGTCGAGGCACACAAAGAAGAGGTGAC 620
           ||   |  |  |  |  |  |  | |||||||  |  || ||  ||   |   || | ||
Qy    1034 TGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCAC 1093

Db     621 AGATCTCATGCTGTGTGCAGGAGAGATGGATGGGGGCAAAGACACTTGCAAGGGTGACTC 680
           ||||  |||| | ||||||||  ||  | |  |||||   ||||| ||| |||| || ||
Qy    1094 AGATGGCATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTC 1150

Db     681 AGGAGGCCCCCTCATCTGTAATGGTGTGCTCCAAGGCATCACGTCCTGGGGCTTTAACCC 740
           ||||||||||  | ||| ||||||  ||||| ||||||| ||||||||||   ||||
Qy    1151 TGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCC 1210

Db     741 ATGCGGTGAACCCAAGAAGCCAGGCATCTACACCAA 776
           || ||     || | || || ||| |||| |||||
Qy    1211 CTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT   10
LOCUS       D78203       1419 bp    mRNA            PRI       28-JAN-1997
DEFINITION  Human mRNA for neurosin, complete cds.
ACCESSION   D78203
NID         g1805492
KEYWORDS    neurosin; trypsin-like serine protease precursor.
SOURCE      Homo sapiens Colon Adenocarcinoma cell_line:COLO 201 cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (sites)
  AUTHORS   Yamashiro,K., Tsuruoka,N., Kodama,S., Tsujimoto,M., Yamamura,Y.,
            Tanaka,T., Nakazato,H. and Yamaguchi,N.
  TITLE     Molecular cloning of a novel trypsin-like serine protease
            (neurosin) preferentially expressed in brain
  JOURNAL   Biochim. Biophys. Acta 1350 (1), 11-14 (1997)
  MEDLINE   97157069
REFERENCE   2  (bases 1 to 1419)
  AUTHORS   Tsuruoka,N., Yamashiro,K., Tsujimoto,M., Yamamura,Y. and
            Yamaguchi,N.
  TITLE     Molecular Cloning of a Novel Trypsin-like Serine Protease (SP59)
            Preferentially Expressed in Brain
  JOURNAL   Unpublished (1996)
REFERENCE   3  (bases 1 to 1419)
  AUTHORS   Tsuruoka,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-1995) to the DDBJ/EMBL/GenBank databases. Nobuo
            Tsuruoka, Suntory Institute for Biomedical Research, Pharmaceutical
            Research Laboratories I; 1-1-1 Wakayamadai, Shimamoto-cho,
            Mishima-gun, Osaka 618, Japan
            (E-mail:tsuruoka@minase.suntory.co.jp, Tel:075-962-9285(ex.3110),
            Fax:075-962-6448)
FEATURES             Location/Qualifiers
     source          1. .1419
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="COLO 201"
                     /cell_type="Adenocarcinoma"
                     /tissue_type="Colon"
     CDS             156. .890
                     /note="trypsin-like serine protease precursor"
                     /codon_start=1
                     /product="neurosin"
                     /db_xref="PID:d1011968"
                     /db_xref="PID:g1805493"
                     /translation="MKKLMVVLSLIAAAWAEEQNKLVHGGPCDKTSHPYQAALYTSGH
                     LLCGGVLIHPLWVLTAAHCKKPNLQVFLGKHNLRQRESSQEQSSVVRAVIHPDYDAAS
                     HDQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWGKTADGDFPDTIQCAYI
```

```
                     HLVSREECEHAYPGQITQNMLCAGDEKYGKDSCQGDSGGPLVCGDHLRGLVSWGNIPC
                     GSKEKPGVYTNVCRYTNWIQKTIQAK"
     mat_peptide     219. .887
                     /product="neurosin"
BASE COUNT      315 a    432 c    362 g    310 t
ORIGIN

  Query Match           5.3%;  Score 71;  DB 21;  Length 1419;
  Best Local Similarity  65.7%;  Pred. No. 9.54e-35;
  Matches   180;  Conservative    0;  Mismatches  91;  Indels   3;  Gaps   3;

Db     604 ATTTCCCTGACACCATCCAGTGTGCATACATCCACCTGGTGTCCCGTGAGGAGTGTGAGC 663
           |||| ||||||||  || | ||||||  |   |   |   |  |||   || ||||||||
Qy    1012 ATTTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGG 1071

Db     664 ATGCCTACCCTGGCCAGATCACCCAGAACATGTTGTGTGCTGGGGATGAGAAGTACGGGA 723
           |||| ||||| || ||||||||  |   |||| | ||||| ||  |  || |  | |||
Qy    1072 ATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGC-AGCAGCAA-AGGGGC 1129

Db     724 AGGATTCCTGCCAGGGTGATTCTGGGGGTCCGCTGGTATGTGGAGACCACCTCCGAGGCC 783
           | |  | |||||||| |||||||| || || ||||| |||| |     ||||  |||
Qy    1130 TG-ACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCA 1188

Db     784 TTGTGTCATGGGGTAACATCCCCTGTGGATCAAAGGAGAAGCCAGGAGTCTACACCAACG 843
           |    || |||||      |||||||||        || || || || ||||| ||||||
Qy    1189 TCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACA 1248

Db     844 TCTGCAGATACACGAACTGGATCCAAAAAACCAT 877
           ||||| | |||  | |||||||| | || | |||
Qy    1249 TCTGCCGCTACCTGGACTGGATCAAGAAGATCAT 1282


RESULT   11
LOCUS       AF013988     1451 bp    mRNA            PRI       12-AUG-1997
DEFINITION  Homo sapiens serine protease mRNA, complete cds.
ACCESSION   AF013988
NID         g2318114
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 1451)
  AUTHORS   Little,S.P., Johnstone,E.M., Dixon,E.P., Norris,F., Buckley,W.,
            Becker,G., Johnson,M.,., Dobbins,J.R., Wyrick,T., Miller,J.R.,
            MacKellar,W., Hepburn,D., Corvalan,J., McClure,D., Liu,X.,
            Stephenson,D. and Clemens,J.
  TITLE     Zyme cDNA isolated from AD brain tissue
  JOURNAL   J. Biol. Chem. (1997) In press
REFERENCE   2  (bases 1 to 1451)
  AUTHORS   Little,S.P., Johnstone,E.M. and Norris,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-1997) CNS Division, Eli Lilly and Company, Lilly
            Corporate Center, Indianapolis, IN 46285, USA
FEATURES             Location/Qualifiers
     source          1. .1451
                     /organism="Homo sapiens"
                     /strain="human"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="19q13.3"
                     /tissue_type="Alzheimer's disease brain tissue"
     CDS             147. .881
                     /note="Zyme; protease bears homology to Kallikrein class
                     and can be localized to microvessels and microglia;
                     chymotrypsin-like"
                     /codon_start=1
                     /product="serine protease"
                     /db_xref="PID:g2318115"
                     /translation="MKKLMVVLSLIAAAWAEEQNKLVHGGPCDKTSHPYQAALYTSGH
                     LLCGGVLIHPLWVLTAAHCKKPNLQVFLGKHNLRQRESSQEQSSVVRAVIHPDYDAAS
                     HDQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWGKTADGDFPDTIQCAYI
```

```
                     HLVSREECEHAYPGQITQNMLCAGDEKYGKDSCQGDSGGPLVCGDHLRGLVSWGNIPC
                     GSKEKPGVYTNVCRYTNWIQKTIQAK"
BASE COUNT      342 a    434 c    367 g    308 t
ORIGIN

  Query Match           5.3%;  Score 71;  DB 21;  Length 1451;
  Best Local Similarity  65.7%;  Pred. No. 9.54e-35;
  Matches   180;  Conservative    0;  Mismatches  91;  Indels   3;  Gaps   3;

Db     595 ATTTCCCTGACACCATCCAGTGTGCATACATCCACCTGGTGTCCCGTGAGGAGTGTGAGC 654
           |||| ||||||||  || | ||||||  |   |   |   |  |||   || ||||||||
Qy    1012 ATTTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGG 1071

Db     655 ATGCCTACCCTGGCCAGATCACCCAGAACATGTTGTGTGCTGGGGATGAGAAGTACGGGA 714
           |||| ||||| || ||||||||  |   |||| | ||||| ||  |  || |  | |||
Qy    1072 ATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGC-AGCAGCAA-AGGGGC 1129

Db     715 AGGATTCCTGCCAGGGTGATTCTGGGGGTCCGCTGGTATGTGGAGACCACCTCCGAGGCC 774
           | |  | |||||||| |||||||| || || ||||| |||| |     ||||  |||
Qy    1130 TG-ACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCA 1188

Db     775 TTGTGTCATGGGGTAACATCCCCTGTGGATCAAAGGAGAAGCCAGGAGTCTACACCAACG 834
           |    || |||||      |||||||||        || || || || ||||| ||||||
Qy    1189 TCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACA 1248

Db     835 TCTGCAGATACACGAACTGGATCCAAAAAACCAT 868
           ||||| | |||  | |||||||| | || | |||
Qy    1249 TCTGCCGCTACCTGGACTGGATCAAGAAGATCAT 1282


RESULT   12
LOCUS       S82666       1457 bp    mRNA            PRI       10-FEB-1997
DEFINITION  NES1=normal epithelial cell specific gene 1 [human, mammary
            epithelial cells, 76N, mRNA, 1457 nt].
ACCESSION   S82666
NID         g1835924
KEYWORDS    .
SOURCE      human 76N mammary epithelial cells.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1457)
  AUTHORS   Liu,X.L., Wazer,D.E., Watanabe,K. and Band,V.
  TITLE     Identification of a novel serine protease-like gene, the expression
            of which is down-regulated during breast cancer progression
  JOURNAL   Cancer Res. 56 (14), 3371-3379 (1996)
  MEDLINE   96320486
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 178916] from the original journal article.
            This sequence comes from Fig. 2.
FEATURES             Location/Qualifiers
     source          1. .1457
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            82. .900
                     /note="normal epithelial cell specific gene 1"
                     /gene="NES1"
     CDS             82. .900
                     /gene="NES1"
                     /note="serine protease homolog; This sequence comes from
                     Fig. 2.  Protein sequence is in conflict with the
                     conceptual translation; Arginine-274 codon lacks initial
                     nucleotide C"
                     /codon_start=1
                     /db_xref="PID:g1835925"
                     /translation="MRAPHLHLSAASGARALAKLLPLLMAQLWAAEAALLPQNDTRLD
                     PEAYGAPCARGSQPWQVSLFNGLSFHCAGVLVDQSWVLTAAHCGNKPLWARVGDDHLL
                     LLQGEQLRRTTRSVVHPKYHQGSGPILPRRTDEHDLMLLKLARPVVPGPRVRALQLPY
                     RCAQPGDQCQVAGWGTTAARRVKYNKGLTCSSITILSPKECEVFYPGVVTNNMICAGL
                     DRGQDPCQSDSGGPLVCDETLQGILSWGVYPCGSAQHPAVYTQICKYMSWINKVIRSN
                     "
BASE COUNT      293 a    480 c    377 g    307 t
ORIGIN
```

```
  Query Match             5.3%;  Score 71;  DB 21;  Length 1457;
  Best Local Similarity  60.7%;  Pred. No. 9.54e-35;
  Matches   201;  Conservative    0;  Mismatches 130;  Indels   0;  Gaps   0;

Db    571 CAGCCCGGAGACCAGTGCCAGGTTGCTGGCTGGGGCACCACGGCCGCCCGGAGAGTGAAG 630
          ||||| ||  |  |||||    ||  | ||||||||||      ||   |   ||| |||
Qy    954 CAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAAT 1013

Db    631 TACAACAAGGGCCTGACCTGCTCCAGCATCACTATCCTGAGCCCTAAAGAGTGTGAGGTC 690
          |       |    || | |||  |     | |  ||| |   ||  ||  |||||||||
Qy   1014 TTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGAT 1073

Db    691 TTCTACCCTGGCGTGGTCACCAACAACATGATATGTGCTGGACTGGACCGGGGCCAGGAC 750
            ||||| ||   | ||||  |   |||| | ||||| ||      |   ||    |||
Qy   1074 GCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGAC 1133

Db    751 CCTTGCCAGAGTGACTCTGGAGGCCCCCTGGTCTGTGACGAGACCCTCCAAGGCATCCTC 810
          | ||||||  | ||  |||||||||||||||||  |||||  |   | |||||  ||||||
Qy   1134 ACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACA 1193

     811 TCGTGGGGTGTTTACCCCTGTGGCTCTGCCCAGCATCCAGCTGTCTACACCCAGATCTGC 870
          || |||||     ||||||||||     || |  | || |  ||||| ||| | ||||||
Qy   1194 TCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACATCTGC 1253

Db    871 AAATACATGTCCTGGATCAATAAAGTCATAG 901
            ||| ||  |||||||||  ||  ||||||
Qy   1254 CGCTACCTGGACTGGATCAAGAAGATCATAG 1284


RESULT   13
LOCUS       HSU62801     1506 bp    mRNA            PRI       02-SEP-1996
DEFINITION  Human protease M mRNA, complete cds.
ACCESSION   U62801
NID         g1518787
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1506)
  AUTHORS   Anisowicz,A., Sotiropoulou,G., Stenman,G., Mok,S.C. and Sager,R.
  TITLE     A Novel Protease Homolog Differentially Expressed in Breast and
            Ovarian Cancer
  JOURNAL   Mol. Med. (Camb. Mass.) (1996) In press
REFERENCE   2  (bases 1 to 1506)
  AUTHORS   Anisowicz,A., Sotiropoulou,G. and Sager,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-1996) Cancer Genetics, Dana-Farber Cancer
            Institute/Harvard Medical School, 44 Binney, Boston, MA 02115, USA
FEATURES             Location/Qualifiers
     source          1. .1506
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="76N normal mammary epithelial cells"
                     /chromosome="19"
                     /map="19q13.3"
     sig_peptide     246. .293
     CDS             246. .980
                     /note="serine protease homolog of the trypsin family"
                     /codon_start=1
                     /product="protease M"
                     /db_xref="PID:g1518788"
                     /translation="MKKLMVVLSLIAAAWAEEQNKLVHGGPCDKTSHPYQAALYTSGH
                     LLCGGVLIHPLWVLTAAHCKKPNLQVFLGKHNLRQRESSQEQSSVVRAVIHPDYDAAS
                     HDQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWGKTADGDFPDTIQCAYI
                     HLVSREECEHAYPGQITQNMLCAGDEKYGKDSCQGDSGGPLVCGDHLRGLVSWGNIPC
                     GSKEKPGVYTNVCRYTNWIQKTIQAK"
     misc_feature    294. .308
                     /note="encodes activation peptide"
     mat_peptide     309. .977
                     /product="protease M"
     misc_feature    429. .431
                     /note="encodes His-62 catalytic triad"
     misc_feature    561. .563
                     /note="encodes Asp-106 catalytic triad"
     misc_structure  645. .653
                     /note="encodes glycosylation motif: AsnThrThr"
     misc_feature    816. .818
                     /note="encodes active site for trysin-like cleavage:
                     Asp-191"
     misc_feature    834. .836
                     /note="encodes Ser-197 catalytic triad"
     polyA_signal    1095. .1100
                     /note="pancreas-specific"
     polyA_signal    1490. .1495
BASE COUNT      321 a    465 c    396 g    324 t
ORIGIN

  Query Match             5.3%;  Score 71;  DB 21;  Length 1506;
  Best Local Similarity  65.7%;  Pred. No. 9.54e-35;
  Matches   180;  Conservative    0;  Mismatches  91;  Indels   3;  Gaps   3;

Db    694 ATTTCCCTGACACCATCCAGTGTGCATACATCCACCTGGTGTCCCGTGAGGAGTGTGAGC 753
          ||||  ||||||||  ||  | ||||||  |   |   |   |  |  |||   || ||||||||
Qy   1012 ATTTTCCTGACACTCTCAACTGTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGG 1071

Db    754 ATGCCTACCCTGGCCAGATCACCCAGAACATGTTGTGTGCTGGGGATGAGAAGTACGGGA 813
          ||||  |||||  ||  |||||||||   |   ||||  |  |||||  ||   |   ||  |  |||
Qy   1072 ATGCTTACCCGGGGCAGATCACAGATGGCATGGTCTGTGCAGGC-AGCAGCAA-AGGGGC 1129

Db    814 AGGATTCCTGCCAGGGTGATTCTGGGGGTCCGCTGGTATGTGGAGACCACCTCCGAGGCC 873
            |  |  | ||||||||  ||||||||  ||  ||  |||||  ||||   |      ||||  |||
Qy   1130 TG-ACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCACTCCAGGGCA 1188

Db    874 TTGTGTCATGGGGTAACATCCCCTGTGGATCAAAGGAGAAGCCAGGAGTCTACACCAACG 933
          |    ||  |||||       |||||||||        ||  ||  ||  ||  |||||  ||||||
Qy   1189 TCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATACCAACA 1248

Db    934 TCTGCAGATACACGAACTGGATCCAAAAAACCAT 967
          |||||  |  |||   |  ||||||||  |  ||  |  |||
Qy   1249 TCTGCCGCTACCTGGACTGGATCAAGAAGATCAT 1282


RESULT   14
LOCUS       RATKALLB     762 bp    mRNA            ROD       03-JAN-1995
DEFINITION  Rattus norvegicus (clone RSKG50) kallikrein mRNA, 3' end.
ACCESSION   L33840
NID         g609584
KEYWORDS    kallikrein.
SOURCE      Rattus norvegicus (strain Sprague-Dawley) (tissue library: RT-PCR)
            submandibular gland cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Zintz,C.B., Ma,J.-X., Chao,J. and Chao,L.
  TITLE     Isolation and characterization of a new rat kallikrein cDNA with
            predominant expression in the kidney
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="submandibular gland"
                     /tissue_lib="RT-PCR"
     CDS             <1. .720
                     /codon_start=1
                     /product="kallikrein"
                     /db_xref="PID:g609585"
                     /translation="AAPPGQSRVVGGFKCEKNSQPWQVAVINEDLCGGVLIDPSWVIT
                     AAHCYSDNYHVLLGQNNLSEDVQHRLVSQSFRHPDYKPFLMRNHTRKPKDYSNDLMLL
                     HLSEPADITDGVKVIDLPTKEPKVGSTCLVSGWGSTNPSEWEFPDDLQCVNIHLLSNE
                     KCIKAYKEKVTDLMLCAGELEGGKDTCRGDSGGPLICDGVLQGITSWGSVPCGEPNKP
```

```
                    GIYTKLIKFTSWIKEVMKKNP"
BASE COUNT      208 a    204 c    189 g    161 t
ORIGIN

  Query Match           5.2%;  Score 70;  DB 22;  Length 762;
  Best Local Similarity  60.9%;  Pred. No. 6.44e-34;
  Matches   238;  Conservative    0;  Mismatches 150;  Indels   3;  Gaps   2;

Db    278 AGGACTACAGCAATGACCTGATGCTGCTCCACCTCAGCGAGCCTGCGGACATCACAGATG 337
          ||||| ||| | |||| |||||||| || || ||  | || |  ||   | |
Qy    859 AGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCCA 918

Db    338 GGGTGAAGGTCATCGATCTGCCCACGAAGGAGCCCAAGGTGGGGAGCACCTGCCTTGTCT 397
            ||||||  ||||   ||| |     |     || ||   ||    |  |||   ||||
Qy    919 AAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCT 978

Db    398 CAGGCTGGGGCAGCACCAACCCCTCTGAGTGGGAATTCCCTGATGATCTCCAGTGTGTGA 457
          ||||||||||||    || |    |       | | || |||||   |||| | ||||
Qy    979 CAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAG 1038

Db    458 ACATCCACCTACTGTCTAATGAGAAGTGCATCAAAGCCTACAAAGAAAAGGTGACAGATC 517
          |  |  |  |  |  |  |  |  |||||||     | || |||   |   || | ||||||
Qy   1039 AAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATG 1098

Db    518 TGATGCTGTGTGCAGGAGAGTTGGAAGGAGGCAAAGACACTTGCAGGGGTGACTCAGGAG 577
            ||| | ||||||||  ||  | || | |||   ||||| |||  ||| || || ||||
Qy   1099 GCATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTCTGGAG 1155

Db    578 GCCCCCTCATCTGTGATGGTGTGCTCCAGGGCATCACATCATGGGGCTCTGTCCCATGCG 637
          |||||||  | ||||||||||  |||||||||||||||| |||||||| | ||| || |
Qy   1156 GCCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTG 1215

Db    638 GTGAACCCAATAAGCCAGGCATCTACACCAA 668
          |     || | || || ||| |||| |||||
Qy   1216 GGAGGTCCGACAAACCTGGCGTCTATACCAA 1246


RESULT   15
LOCUS       RATKALSA      518 bp    mRNA            ROD       01-SEP-1988
DEFINITION  Rat submaxillary gland S1 kallikrein mRNA, partial cds.
ACCESSION   M11564
NID         g205031
KEYWORDS    .
SOURCE      Rat submaxillary gland, cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Rattus.
  REFERENCE 1  (bases 1 to 518)
  AUTHORS   Ashley,P.L. and MacDonald,R.J.
  TITLE     Kallikrein-related mRNAs of the rat submaxillary gland: Nucleotide
            sequences of four distinct types including tonin
  JOURNAL   Biochemistry 24, 4512-4520 (1985)
  MEDLINE   86051477
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             <1. .473
                     /note="S1 kallikrein"
                     /codon_start=3
                     /db_xref="PID:g205032"
                     /translation="LIWNHTRKPGDYSNDLMLLHLSEPADITDGVKVIDLPTKEPKVG
                     STCLVSGWGSTNPSEWEFPDDLQCVNIHLLSNEKCIKAYKEKVTDLMLCAGELEGGKD
                     TCRGDSGGPLICDGVLQGITSWGSVPCGEPNKPGIYTKLIKFTSWIKEVMKENP"
BASE COUNT      144 a    141 c    132 g    101 t
ORIGIN

  Query Match           5.1%;  Score 69;  DB 22;  Length 518;
  Best Local Similarity  60.8%;  Pred. No. 4.33e-33;
  Matches   237;  Conservative    0;  Mismatches 150;  Indels   3;  Gaps   2;

Db     32 GGACTACAGCAATGACCTGATGCTGCTCCACCTCAGCGAGCCTGCGGACATCACAGATGG 91
          |||| ||| | |||| |||||||| || || ||  | || |  ||   | |
Qy    860 GGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCCAA 919

Db     92 GGTGAAGGTCATCGATCTGCCCACGAAGGAGCCCAAGGTGGGGAGCACCTGCCTTGTCTC 151
            ||||||  ||||   ||| |     |     || ||   ||    |  |||   |||||
Qy    920 AGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTC 979

Db    152 AGGCTGGGGCAGCACCAACCCCTCTGAGTGGGAATTCCCTGATGATCTCCAGTGTGTGAA 211
          ||||||||||||    || |    |       | | || |||||   |||| | ||||   |
Qy    980 AGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAGA 1039

Db    212 CATCCACCTACTGTCTAATGAGAAGTGCATCAAAGCCTACAAAGAAAAGGTGACAGATCT 271
            |  |  |  |  |  |  |  |  |||||||     | || |||   |   || | ||||||
Qy   1040 AGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGG 1099

Db    272 GATGCTGTGTGCAGGAGAGTTGGAAGGGGGCAAAGACACTTGCAGGGGTGACTCAGGAGG 331
            ||| | ||||||||  ||  | || |||||   ||||| |||  ||| || || |||||
Qy   1100 CATGGTCTGTGCAGGC-AGCAGCAAAGGGGCT--GACACGTGCCAGGGCGATTCTGGAGG 1156

Db    332 CCCCCTCATCTGTGATGGTGTGCTCCAAGGCATCACATCATGGGGCTCTGTCCCATGCGG 391
          ||||||  | ||||||||||  ||||| ||||||||||| |||||||| | ||| || ||
Qy   1157 CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGG 1216

Db    392 TGAACCCAATAAGCCAGGCATCTACACCAA 421
            || | || || ||| |||| |||||
Qy   1217 GAGGTCCGACAAACCTGGCGTCTATACCAA 1246


Search completed: Mon Aug  3 11:36:24 1998
Job time : 1759 secs.
```